(12) **UK Patent Application** (19) **GB** (11) **2 226 675** (13) **A**

(43) Date of A publication 04.07.1990

(21) Application No 8929267.6

(22) Date of filing 28.12.1989

(30) Priority data
(31) 63330905 (32) 29.12.1988 (33) JP

(71) Applicant
Kabushiki Kaisha Toshiba

(Incorporated in Japan)

72, Horikawa-Cho, Saiwai-Ku, Kawasaki-Shi, Kanagawa-Ken, Japan

(72) Inventor
Yasuyuki Masai

(74) Agent and/or Address for Service
Batchellor Kirk & Eyles
2 Pear Tree Court, Farringdon Road, London, EC1R 0DS, United Kingdom

(51) INT CL⁵
G10L 3/00, H04M 1/26

(52) UK CL (Edition K)
G4R RHA R1F R10A R10X R11D R11E R11F R12F
R3A R3G R3X R9B R9C
H4K KBNJ KFH
U1S S2215

(56) Documents cited
GB 1055371 A GB 0974850 A

(58) Field of search
UK CL (Edition J) G4R REX RHA RHB
INT CL⁴ G10L
Online databases: WPI, INSPEC

(54) **Voice recognition apparatus**

(57) Voice recognition apparatus recognises words, e.g. names, by comparing an input voice with previously registered voice pattern data (3). A voice recorder (4) records the spoken words associated with the registered voice pattern data, a voice reproducer (7) reproducing the recorded words in response to a command. A manually-actuated voice designator (8) selects a desired word when the desired word is reproduced. The words may be names associated with telephone numbers to be dialled. Operation of the voice designator (8) deletes the selected word from the voice pattern data (3) and the voice recorder (4).

In an alternative embodiment recognition of a spoken word causes a confirmation word or phase to be read from the voice recorder; the confirmation word or phase may be the recognised word or be independent thereof.
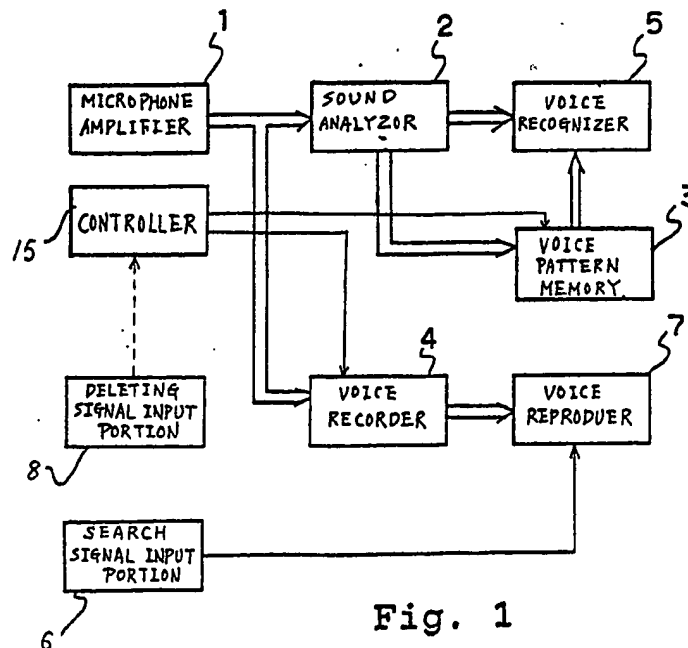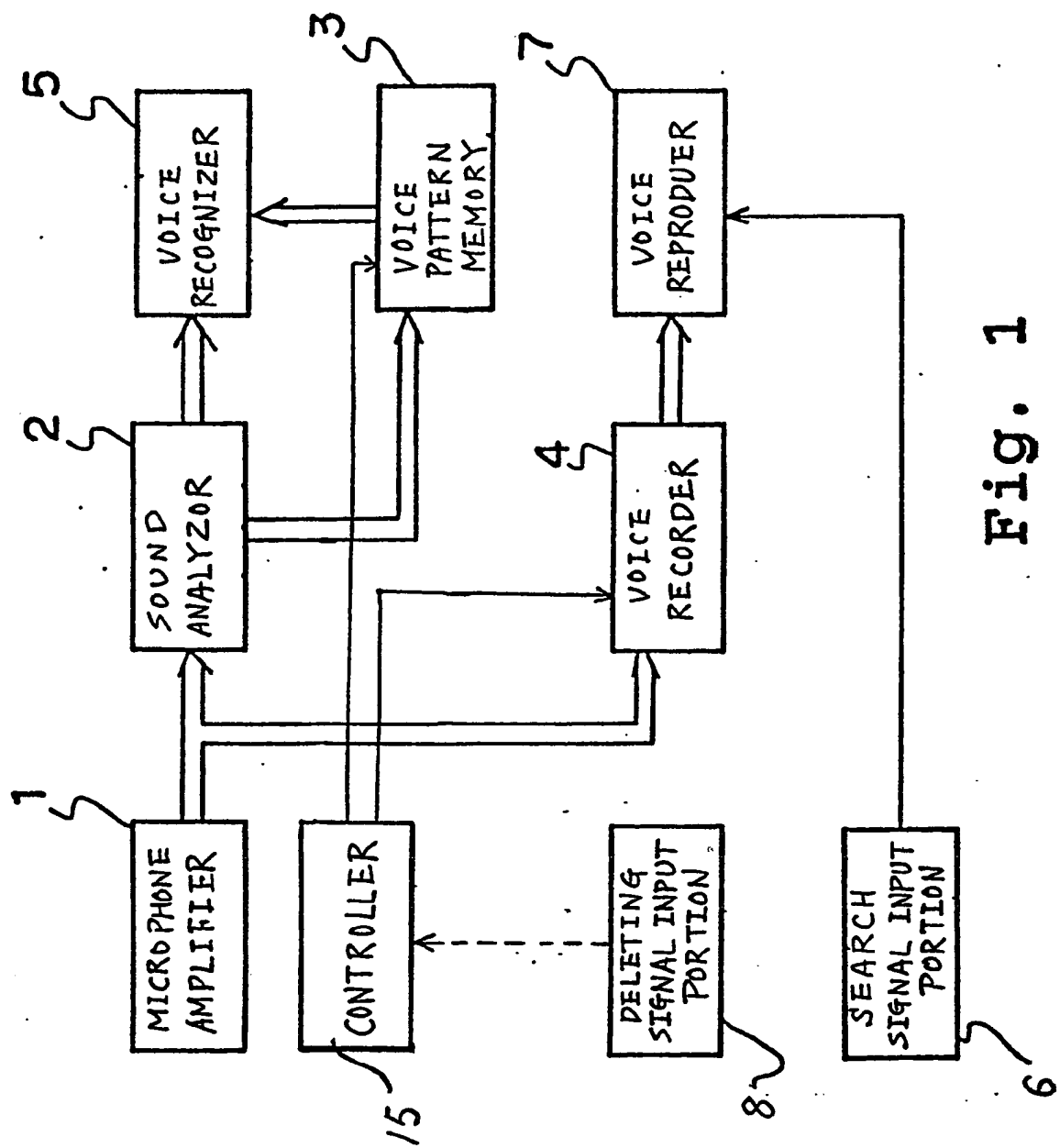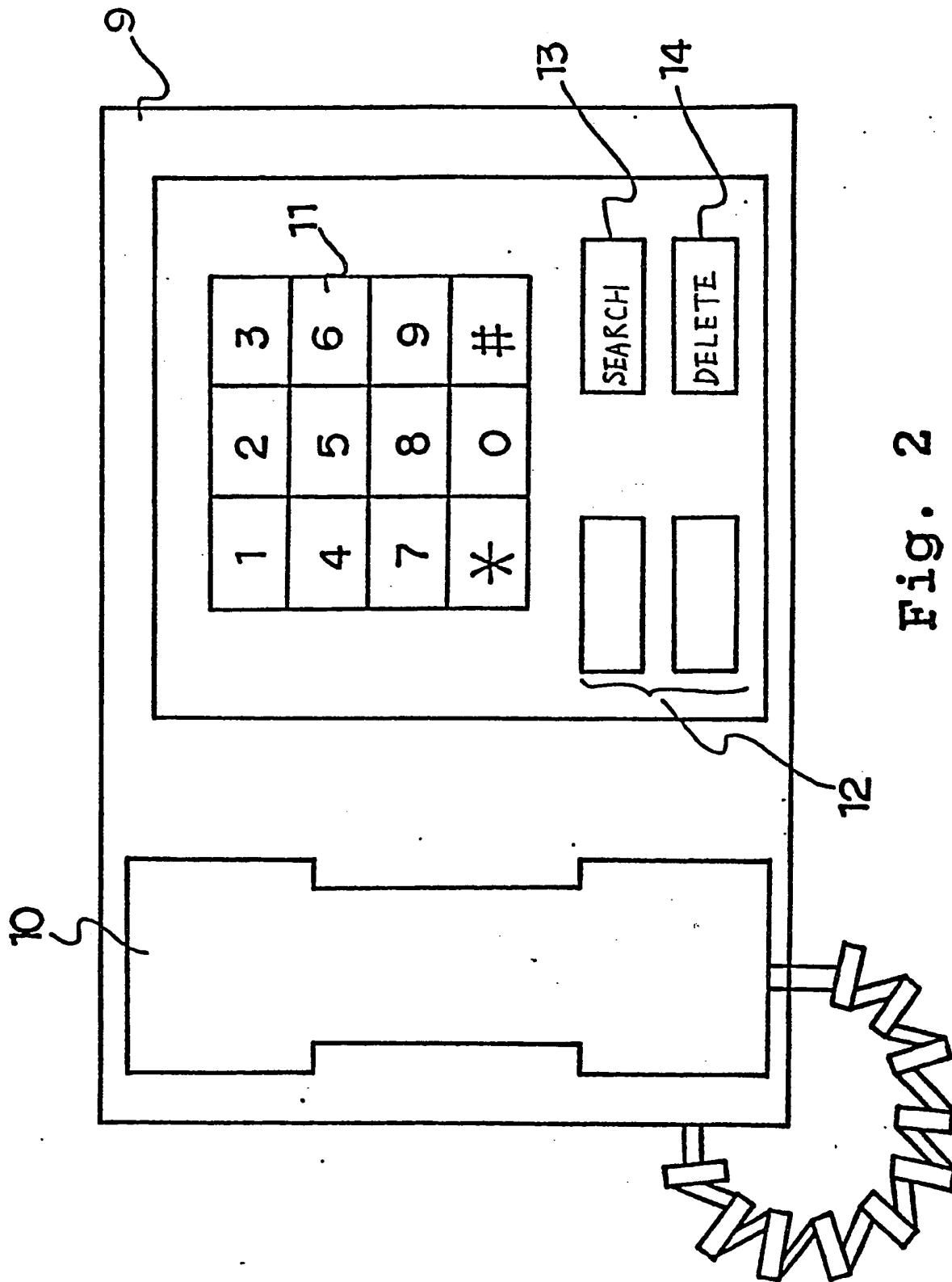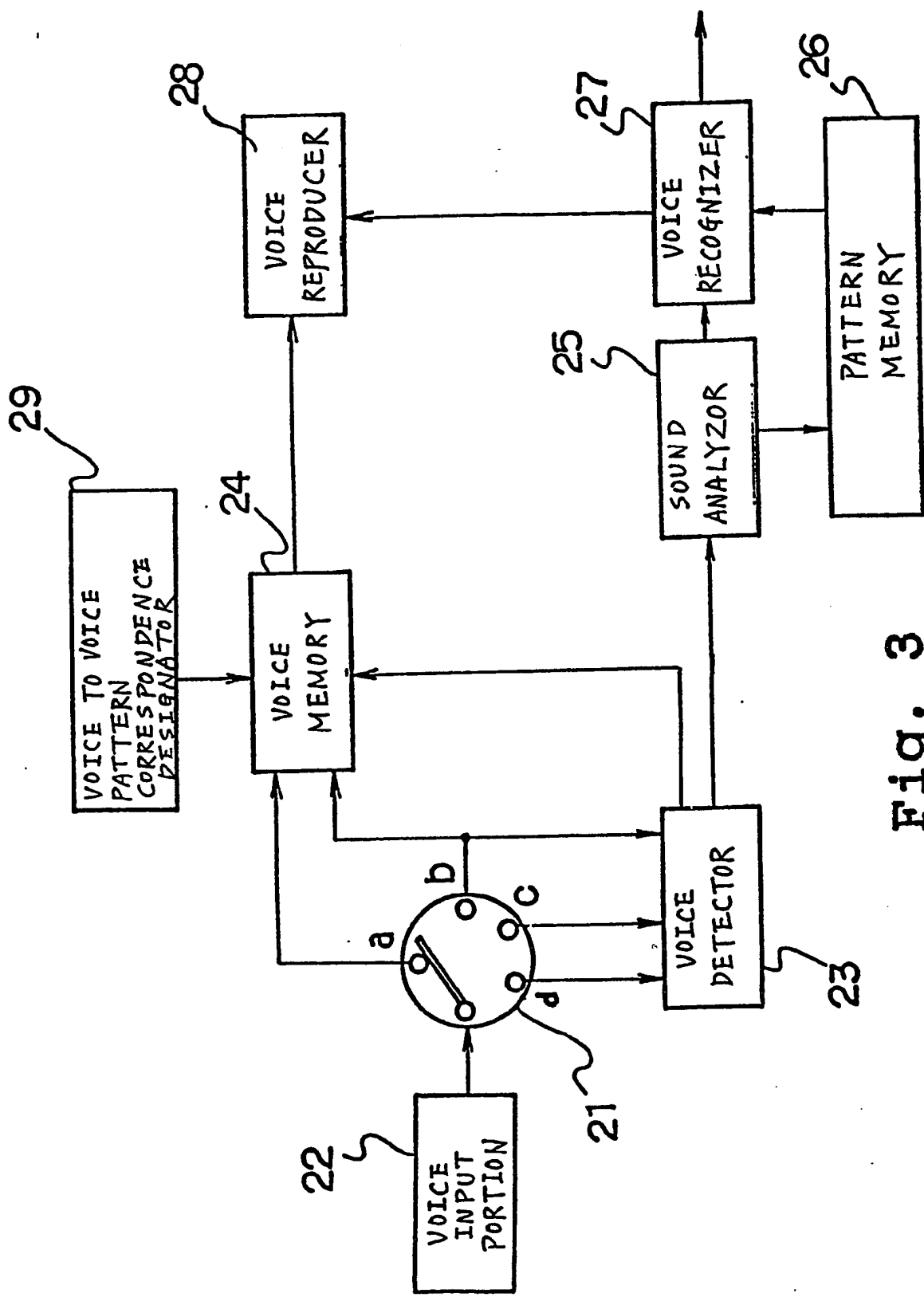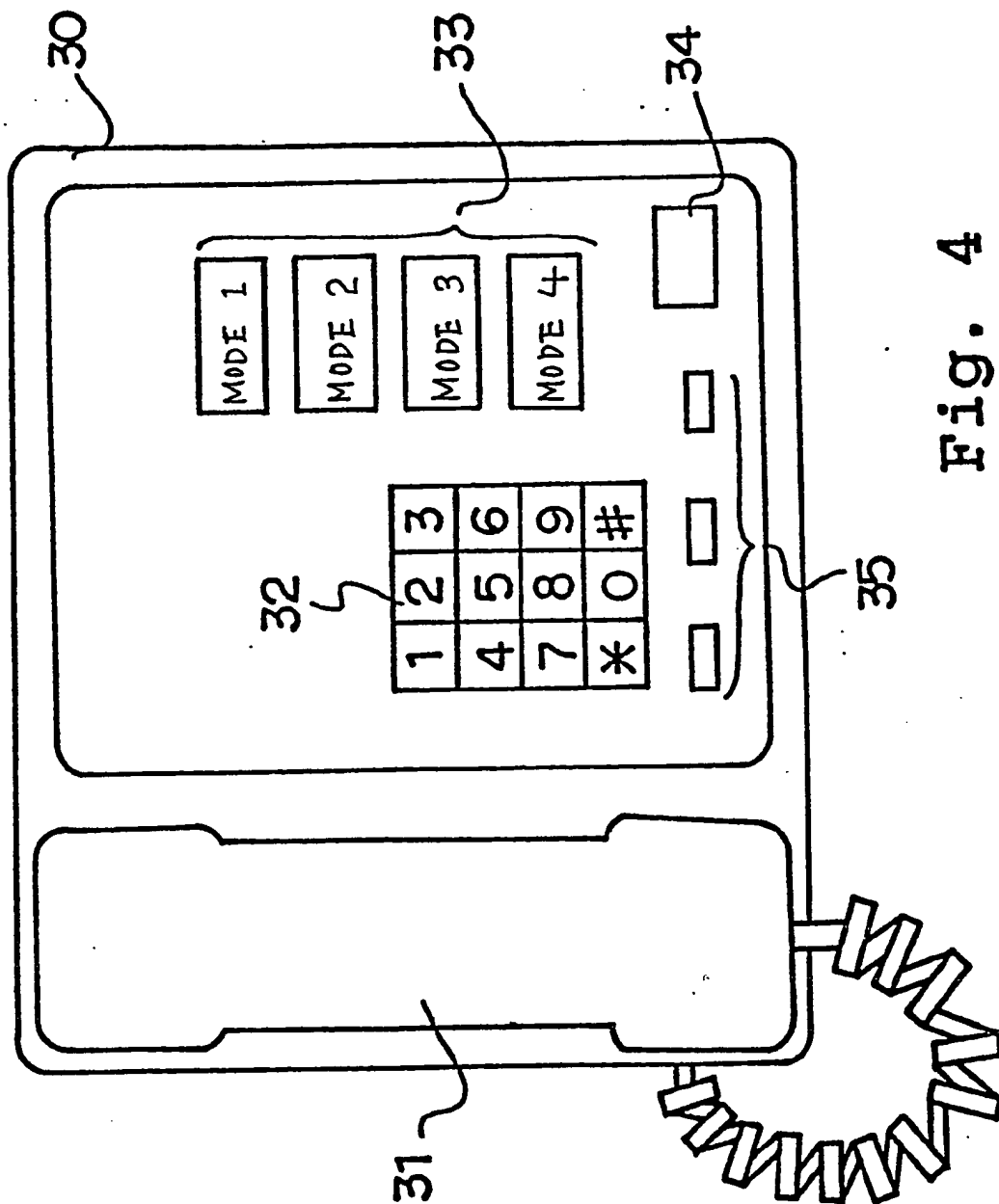
Fig. 1

Fig. 1

Fig. 2

Fig. 3

Fig. 4

VOICE RECOGNITION APPARATUS AND

TELEPHONE SET EQUIPPED WITH THE APPARATUS

The present invention relates to a voice recognition apparatus and, more particularly, to a telephone set equipped with such a voice recognition apparatus.

In recent years, a telephone set which dials a telephone number corresponding to the name of a subscriber, which has been stored in advance, according to the result of voice recognition of a person who spoke the name, has been put in practical use.

Now, with such a telephone set, to delete a name once stored or to re-store a name, it is necessary to designate the name to be deleted or re-stored.

As a method used so far to designate this name, it is general practice to decide names and corresponding telephone numbers in advance, designate a desired name by inputting the corresponding number or by storing names in advance together with voice patterns and characters of the names and designating a desired name by indicating the corresponding character on an indicator.

However, in a method to designate a name by designating a number, as described above, it is necessary for a user always to keep the relationship between name and number in mind. Further, in a method

to store a name with a voice pattern and characters, cost becomes high as a display is needed for indicating characters and, furthermore, there is a problem that it is troublesome to input characters.

As described above, an existing telephone set equipped with a voice recognition apparatus has such problems that the user is required always to keep the correspondence between names and numbers in order to designate an already stored name for some purpose. Further, a display is required to indicate the name, and it is necessary to input characters to store the name.

The present invention, therefore, seeks to provide a new and improved voice recognition apparatus and telephone set equipped with the apparatus in order to solve such problems as mentioned above.

A voice recognition apparatus comprises means for recording input voices applied to the apparatus; means for producing and storing voice pattern data of each of the input voices; means for comparing voice pattern data of a voice applied to the apparatus with voice pattern data stored therein; means for reproducing the recorded input voices on command; means for selecting one of the reproduced input voices; and means for processing the selected input voice and the recorded voice pattern data of that voice in a prescribed manner.

In the first embodiment of the present invention, voices recorded by the voice recorder are reproduced by the voice reproducer as required and, when a desired voice is reproduced, that voice is designated by a manual voice operation of a designator. Then, a prescribed processing pre-defined in this apparatus is carried out for the designated voice and the voice pattern data of the voice.

A third embodiment of the present invention is directed to a telephone set for automatically dialling a prescribed telephone number in response to a voice pattern of a calling voice corresponding to a person name to be called by comparing the voice pattern of the calling voice with reference to previously registered voice pattern data indicating the personal name which is previously registered therein. The telephone set includes a voice recorder for recording a voice associated with the registered voice pattern data of the voices, a voice reproducer for reproducing the voice recorded in the voice recorder in response to a command, a voice designator for designating a desired voice when the desired voice has been reproduced and a processor for deleting the designated recorded voice and the voice pattern data registered in correspondence to the recorded voice to the voice designator.

Further in the third embodiment of the present invention, the voice recorded by the voice

recorder is reproduced by the voice reproducer as required and, when a desired voice is reproduced, that voice is designated by the voice designator. Then, the designated voice and its associated voice pattern data are automatically deleted by the processor.

According to the above embodiments of the present invention, therefore, only when a desired voice is designated upon hearing voices being reproduced, such predetermined processings as deletion of the designated voice corresponding to a personal name and its voice pattern, etc., are carried out and the burden on the user can be significantly reduced.

A fourth embodiment of the present invention is directed to a voice recognition apparatus for recognising voice by comparing a voice with reference to a standard voice data which is previously produced from voice patterns corresponding to voices with a same meaning spoken by a plurality of persons and registered therein and outputting a confirmation voice corresponding to the standard voice pattern data. The apparatus includes a voice recorder for recording the voice associated with the registered voice pattern data as a confirmation voice to be output corresponding to the result of recognition, and a switch for turning ON or OFF the recording of the verbal confirmation as occasion demands.

In the fourth embodiment of the present

invention, it is possible to optionally set a person whose voice is to be used as a confirtmation voice for the result of recognition as the recording of a confirmation voice by the voice recorder which can be turned ON/OFF as occasion demands.

A fifth embodiment of the present invention is directed to a voice recognition apparatus for recognising voice by comparing a voice with a reference to a standard voice pattern data which is previously generated from voices corresponding to voices with a same meaning spoken by a plurality of persons and registered therein and outputting a confirmation voice corresponding to the standard voice pattern data. The apparatus includes a voice recorder for recording a confirmation voice to be output corresponding to the result of recognition independently to the registered standard voice pattern data according to the input voices.

In the first embodiment of the present invention, a confirmation voice to be output and corresponding to the result of recognition is recorded by the voice recorder independently to the registered standard voice pattern data by the input of voice and, therefore, the same effect as in the fourth embodiment is obtained.

In order that the invention may be more readily understood, it will now be described, by way of

example only, with reference to the accompanying drawings, in which:-

Fig. 1 is a block diagram showing the configuration of the voice recognition apparatus of a first embodiment of the present invention;

Fig. 2 is a plan showing the external view of a telephone set equipped with the voice recognition apparatus shown in Fig. 1;

Fig. 3 is a block diagram showing the voice recognition apparatus related to another embodiment of the present invention; and

Fig. 4 is a plan showing the external view of a telephone set equipped with the voice recognition apparatus showing in Fig. 3.

Referring now to the drawings, wherein like reference numerals designate identical or corresponding parts throughout the several views, and more particularly to Fig. 1 shows an arrangement of the voice recognition apparatus according to the present invention, which is applied to a telephone set for automatically dialing a telephone number corresponding to a result of recognition of a voice by which personal names, corporation names, etc. desired to be called are input.

As shown in Fig. 1 a voice signal designating a name is input to the telephone set, and analyzed so that a corresponding voice pattern is stored together with the voice signal. More particularly, voice spoken by a user is amplified by a voice input part 1, such as a microphone amplifier, and input into a sound analyzer 2. In the sound analyzer 2, such sound analyses as the band-pass filter

analysis, cepstrum analysis, etc. are carried out as required in order to extract the input voice characteristics. The voice patterns of input voices thus extracted are stored in a voice pattern memory 3. Further, in parallel with this operation, the output of the voice input part 1 is recorded in a voice recorder 4.

Thereafter, a voice spoken is input to the sound analyzer 2 through the voice input part 1 and the sound analysis is carried out. Then, the input voice pattern which is the result of analysis is input to the voice recognizer 5. The voice recognizer 5 performs a matching process between the input voice pattern and the voice pattern stored in the voice pattern memory 3. This matching process may be performed, for instance, using a dynamic programming, partial space method, etc. which are often used, but in this embodiment, any matching processing method may be used.

Then, on the basis of the result of recognition by the voice recognizer 5, a telephone number stored in corresponding with the judged name in the voice pattern storage is searched, and dialing of this telephone set is made using the searched telephone number.

Thus, the processes of name storage by voice and voice recognition are carried out.

Now, on such a telephone set as mentioned above, it often becomes necessary to deletes once stored voice patterns from the voice pattern memory 3.

This deleting method is explained in the following discussion. To delete the already stored specific name "Suzuki", a search signal is output from the search signal

input portion 6 to a voice reproducer 7. When receiving the search signal, the voice reproducer 7 successively reproduces all voices recorded in the voice recorder 4. While listening to names being reproduced by the voice reproducer 7, when "Suzuki" is reproduced, a user inputs a deleting signal from the deleting signal input portion 8 to a controller 15, such as a microcomputer which is generally used for system controls of the apparatus. The controller 15 controls the voice pattern memory 3 and the voice recorder 4 in response to the deleting signal to delete the voice pattern and the stored voice signal of "Suzuki".

Therefore, if a name "Suzuki" was stored for two users (A and B) and user A desired to delete "Suzuki" he stored, if the user A manually inputs characters spelling "Suzuki", while according to this embodiment, it is possible to distinguish "Suzuki" stored by the user B based on the difference in voice of users A and B.

Next, use of the embodiment described in relation to Fig. 1 will be explained below by taking a definite example.

Fig. 2 shows a telephone set equipped with the voice recognition apparatus.

In Fig. 2, numeral 9 designates a telephone casing, 10 a handset, 11 a dial keyboard provided on the surface of the telephone casing 9, and 12 function selector keys which are used for various telephone services.

A search signal input portion 6 and the deleting signal input portion 8 function by an operation of search key 13 and the deleting key 14, respectively, provided at the lower right of the dial keyboard 11 on the surface panel of the

telephone casing 9.

Now, it is assumed that two names "Suzuki" and "Satoh" are both stored by the users A and B, respectively on the telephone set, and it is desired to delete the voice pattern of the name "Satoh" stored by the user A.

In this case, when the search key 13 is first pushed, the voices recorded in the voice recorder 4 are reproduced and heard from the handset 10 in order of "Suzuki" and "Satoh" stored by the user A, and "Suzuki" and "Satoh" stored by the user B. Here, as the name to be deleted is "Satoh" stored by the user A, if the deleting button is pushed when the "Satoh" by the user A is reproduced, "Satoh" stored by the user A is deleted from the voice pattern memory 3 and the voice recorder 4.

Thus, according to this embodiment, a desired name can be be easily designated and the recorded voice stating the name and the the voice pattern of this voice can be easily deleted.

Further, a case where all recorded names are reproduced in order of their storage has been explained in the above embodiment, but a range of names to be reproduced may be restricted or the order of names to be reproduced may be specified.

Though a case where only a specific name among already stored multiple names is designated and deleted has been explained above, the present invention is not limited to this case. The present invention also can be used to realize, for instance, a re-storage function or a function requiring designation of specific name.

Next, another embodiment of the present invention will

be explained hereinafter, in relation to Fig. 3 which is a
block diagram showing another arrangement of the voice
recognition apparatus according to this invention.

In Fig. 3, a switch 21 selects four mode states of the
voice recognition apparatus in this embodiment.

When a contact b of the switch 21 is closed, the input
voice pattern is stored and the voice from the which the
voice pattern is derived is recorded at the same time (Mode
2). When a contact c is closed, the input voice pattern
only is stored and the corresponding voice is not recorded
(Mode 3). When a contact a is closed, no voice pattern is
recorded (Mode 1). When a contact d is closed, the voice
recognition state results (Mode 4).

Next, the state where the contact b of the switch 21 is
closed will be explained in detail.

When a voice is input through the voice input portion
22, that input voice is applied to a voice section detector
23 and at the same time, it is stored or recorded in a voice
memory 24.

The voice section detector 23 is set a threshold value
for the level of the input voice. The voice section
detector 23 detects a voice section by discriminating the
voice section from the absence of sound and noises according
to the threshold value.

The voice memory 24 selectively stores only the voice
signal of the voice section, based on data on a voice
section detected by the voice section detector 23. Then,
this extracted voice signal of the voice section is stored
in the voice memory 24 under control of the detector 23, in
correspondence with the voice pattern obtained by the sound

analyzer 25.

More particularly, the detector 23 also gates the detected voice section to the sound analyzer 25. The sound analyzer 25 analyzes a voice characteristic parameter sequence by filtering the detected voice section is stored in the pattern memory 26 as a standard voice pattern.

After storing the standard voice pattern in the voice pattern memory 26 as described above, and Mode 4 is selected at the time of voice recognition, the voice section is detected by the voice section detector 23 for voices which are input through the voice input portion 22. Similarity of the voice pattern obtained by the sound analyzer 25 for the detected voice section with the standard pattern stored in the voice pattern memory 26 is calculated by the voice recognizer 27, and the result of recognition for the input voice is obtained by mutually comparing the similarity values.

As the the voice recognizing process method used in the voice recognizer 27, the voice recognizing algorithm which has been so far proposed in addition to the similarity calculation may be adopted as appropriate.

A voice signal in the voice storage 24 corresponding to the result of recognition obtained as mentioned above is given to the voice reproducer 28, and a voice for confirming the result of recognition is reproduced and output.

Next, a case where the contact c of the switch 21 is closed will be explained in detail.

In this case, a voice input through the voice input portion 22 is given to the voice detector 23 only and is not given to the voice memory 24, and therefore, no voice signal

is stored in the voice section detector 24. The voice
applied to the voice section detector 23 is processed by the
same manner as in the case where the contact b of the switch
21 is closed, and the voice pattern is stored in the voice
pattern memory 26.

Lastly, a case where the contact a is closed will be
explained in detail.

In this case, a voice input through the voice input
portion 22 is applied only to the voice memory 24 and not to
the voice section detector 23 and therefore, no voice
pattern is stored in the voice pattern memory 26.

The voice memory 24 stores all voices (including noise,
no-sound section, etc.) input through the voice input
portion 22 for the entire period of the voice storage
operation without reference to detection of the input voice
section. As no process for detecting the voice section is
performed, it becomes possible to output voices including a
no-sound section like a sentence as a response to confirm
the result of voice recognition. Here, by a voice to voice
pattern correspondence designator 29, it is designated to
which voice pattern stored in the voice pattern memory 26 a
voice stored in the voice memory 24 corresponds.

Thus, by designating the correspondence of the voice to
the voice pattern, it is possible-to output a verbal
confirmation other than "word" as a response of the result
of recognition.

Fig. 4 shows a telephone set equipped with the voice
recognition apparatus shown in Fig. 3.

In Fig. 4, numeral 30 designates a telephone casing, 31
a handset, and 32 a dial keyboard provided on the surface of

the telephone casing 30. Further, numeral 33 designates a
mode selector switch, including four pushbuttons
corresponding to the switch 21 in Fig. 3. These four
pushbutton switches are structured such that more than two
buttons cannot be pushed simultaneously. Numeral 34
designates a key for designating the correspondence of voice
with the voice pattern together with the dial keyboard 32,
and 35 designates a function selector key used for various
telephone services.

Next explained is a case where two users (A and B) each
store a voice pattern of a voice pronouncing a personal name
"Tanaka" in the voice recognition apparatus, and the
apparatus produces a response as a result of recognition by
means of the voice of "Tanaka" registered by the user.

When the user A is to register his voice pattern, he
sets the mode selector switch 33 in Mode 2, speaks "Tanaka"
and registers his voice pattern and the voice. Then, assume
that the user B is to register his voice pattern, and he
sets the mode selector switch 33 in Mode 3, speaks "Tanaka"
and tries to register his voice pattern. At this time, the
voice "Tanaka" spoken by the user B will not be registered.

Thus, in this example, when the voice patterns by two
users are registered, the verbal confirmation for the result
of recognition is always output by the voice of the user A
whose voice and voice pattern were entered and stored in
Mode 2.

Further, if the user B registers his voice pattern in
Mode 2, the voice of the user B is overwritten to the voice
of the user A and the verbal confirmation for the result of
recognition is output by the voice of the user B.

When it is desired to provide a confirmation response to the users A and B in the voice of another user C, it is necessary to select Mode 1 and register the voice of the user C.

Thus, according to the voice recognition apparatus in this embodiment, whether a voice spoken when storing the voice pattern is to be recorded can be set by the mode selector switch 33, and whose voice is to be used to output the verbal confirmation for the result of recognition can also be decided optionally. In addition, it is possible to return a response by a "word" other than a "word" desired to be recognized by recording a confirming response voice independently of storage of a voice pattern.

As described above, according to the present invention, only by designating a desired name where it is reproduced by hearing a voice spoken to that name, it is possible to perform a prescribed processing such as deletion of a voice corresponding to the name and the voice patterns, etc. and to reduce significantly a burden on users.

In addition, it is also possible to set and output any voice as a confirming response voice for the result of recognition, as desired.

As described above, the present invention can provide an extremely preferable voice recognition apparatus and a telephone set equipped with the apparatus.

Claims:

1.   A voice recognition apparatus comprising
means (4) for recording input voices applied to the
apparatus; means (2, 3) for producing and storing
voice pattern data of each of the input voices; means
(5) for comparing voice pattern data of a voice applied
to the apparatus with voice pattern data stored
therein; means (7) for reproducing the recorded input
voices on command; means (8) for selecting one of the
reproduced input voices; and means (15) for processing
the selected input voice and the recorded voice pattern
data of that voice in a prescribed manner.

2.   Voice recognition apparatus as claimed in
claim 1, wherein the processing means comprise means
whereby the selected input voice and the recorded voice
pattern of that voice are deleted from the recording
means and the voice pattern data store, respectively.

3.   Voice recognition apparatus for recognising
voice by comparing a voice with reference to a
standard voice pattern data which is previously
produced from voice pattern data corresponding to
voices with a same meaning spoken by a plurality of
persons and registered therein and outputting a
confirmation voice corresponding to the standard voice

data, wherein the apparatus comprises voice recorder means (24) for recording a voice associated with the registered voice pattern data as a confirmation voice to be output corresponding to the result of recognition the standard voice pattern by input of the voices; and switching means (21) for turning ON or OFF the recording of the verbal confirmation as occasion demands.

4.   Voice recognition apparatus for recognising voice by comparing a voice with reference to a standard voice pattern data which is previously generated from voices corresponding to voices with a same meaning spoken by a plurality of persons and registered therein and outputting a confirmation voice corresponding to the standard voice pattern data, wherein the apparatus comprises voice recorder means (24, 26) for recording a confirmation voice to be output corresponding to the result of recognition independently to the standard voice pattern data according to the input voice.

5.   Voice recognition apparatus as claimed in any preceding claim, wherein the apparatus is in combination with a telephone set which is arranged to dial a number relating to a voice pattern data which corresponds to voice pattern data stored in the apparatus.